

What is claimed is:

1. A method for real-time online search processing of selected types of information in inter-connected computer networks, the method comprises the steps of:
  - 5 a. assembling site descriptions for a plurality of sites in the inter-connected computer networks including for each of the plurality of sites:
    - i. a URL for the site;
    - ii. a search form URL for the site;
    - iii. generalized rules of how the selected types of information on the site are organized;
    - 10 iv. sample data retrieved from the site corresponding to the selected types of information; and
    - v. descriptions of domains found in the site;
  - b. receiving a request for specified types of information from an online user;
  - c. identifying from the site descriptions, sites which may have the specified types of information;
  - 15 d. constructing search requests for the specified types of information using the site descriptions for each of the identified sites;
  - e. submitting the constructed search requests to the identified sites;
  - f. receiving search results from the identified sites, and upon locating accurate matches in the received search results, extracting information corresponding to the specified types of information in a native language of the site, and displaying the extracted information to the user.
- 20 2. The method of claim 1 wherein the generalized rules include identifying characteristics that can accurately identify the occurrence of the selected types of information within the site.
- 25 3. The method of claim 1 wherein the submitting constructed search requests step is multi-threaded.
4. A method for assembling site information for use in real-time online search processing of selected types of information over inter-connected computer networks, the method comprises the steps of:
  - 30 a. collecting information from a plurality of sites over inter-connected computer networks including:

100-00000000000000000000000000000000

- i. a URL for each of the plurality of sites;
    - ii. a search form URL for each of the plurality of sites;
    - iii. sample pages containing the selected types of information retrieved from each of the plurality of sites;
    - iv. positional information associated with the selected types of information within the sample pages; and
  - b. from the collected information, deriving generalized rules about how the selected types of information on each of the plurality of sites are organized.
- 5 5. The method of claim 4 wherein the deriving generalized rules step
- 10 includes the step of assembling identifying characteristics which can uniquely identify the occurrence of the selected types of information within each of the plurality of sites.
- 15 6. The method of claim 5 wherein the assembling identifying characteristics step includes the steps of
  - i. identifying delimiter characters which bound the selected types of information in the sample pages;
  - ii. retrieving further sample pages from the site; and
  - iii. corroborating the identified delimiter characters against the retrieved further sample pages.
- 20 7. The method of claim 6, wherein the identifying delimiter characters step includes the steps
  - i. from the sample pages, compiling a list of possible strings of delimiter characters for each of the selected types of information.
  - ii. comparing the list of strings of delimiter characters compiled for one of the sample pages with the list of strings of delimiter characters compiled for another of the sample pages, and revising the compiled lists of strings to be consistent for all compared sample pages; and
- 25 further wherein the corroborating step comprises the step of repeating the comparing step for the further sample pages retrieved contemporaneously from the site.
- 30 8. A method for real-time online search processing over inter-connected computer networks, the method comprises the steps of:
  - a. accessing an offline database having vendor descriptions for a plurality of vendor sites over inter-connected computer networks, the vendor descriptions having information about each of the plurality of vendor sites include:

(b) (4) TRADE SECRET

- i. a URL for each of the plurality of vendor sites;
- ii. a search form URL for each of the plurality of vendors;
- iii. description of domains found in each of the plurality of vendor sites;
- 5 iv. generalized rules about how product information is organized on each of the plurality of vendor sites;
- vi. samples of price and product information retrieved from the plurality of vendors;
- b. receiving from an online user a price comparison request for a desired product;
- 10 c. identifying from the vendor descriptions, vendor sites which may have price information relevant to the price comparison request;
- d. constructing search requests for the desired product using the vendor descriptions for each of the identified vendor sites, including the corresponding search form URL;
- 15 e. submitting to the identified vendor sites the constructed search requests;
- f. extracting price and product information from search results received in response to the submitted search requests, wherein the extracted price and product information are in a native language of the site; and
- 20 g. displaying the extracted price and product information to the user.

9. The method of claim 8 wherein the general rules include delimiters which can uniquely identify the occurrence of price and associated information within each of the plurality of vendors.

10. The method of claim 8 wherein the URLs of the plurality of vendor sites, and the extracted price and product information are stored in a database.

25 11. The method of claim 10 wherein the URLs of the plurality of vendor sites, and the extracted price and product information which are stored in the database are updated periodically.

12. The method of claim 11 wherein the URLs of the plurality of vendor sites, and the extracted price and product information, which are stored in the database, are automatically updated daily.

30 13. The method of claim 8 wherein each of the plurality of vendor descriptions is specific to a different online store.

14. The method of claim 13 wherein only one vendor description for each different online store is stored in a database.

15. The method of claim 8 wherein the extracting step includes the step of verifying accurate matches in the search results received in response to the submitted search requests with the desired product.

5 16. The method of claim 8 wherein the displaying step includes the step of displaying price and product information for the desired product only from the vendor site having the best price.

10 17. The method of claim 8 wherein the displaying step includes the step of displaying price and product information for the desired product in a selectable arrangement.

15 18. The method of claim 8 wherein the displaying step includes the step of displaying price and product information for the desired product and which has been sorted according to price.

15 19. The method of claim 8, further includes a step of user authentication comprising a security interface.

20 20. The method of claim 19, wherein the security interface categorizes users as temporary trial and life members.

20 21. The method of claim 8, wherein vendors in the plurality of vendor sites are registered or non-registered vendors.

22. The method of claim 8, wherein the vendor descriptions are automatically constructed through an inductive learning method.

22. The method of claim 22, wherein the inductive learning method can work in multilingual environments.

25 24. The method of claim 22, wherein the inductive learning method is domain independent.

25 25. The method of claim 22, wherein the inductive learning method operates in multiple domains such as books, electronic products, movies, or other products.

30 26. The method of claim 22, wherein the inductive learning method uses a small set of training data.

27. The method of claim 26, wherein the training data includes product examples and the URL from online stores.

28. The method of claim 26, wherein the inductive learning method can extract and identify data independent of presentation style of the online store.

29. A method for real-time online search processing over inter-connected computer networks, the method comprises the steps of:

- a. maintaining in an offline database information for a plurality of vendor sites over inter-connected computer networks, the information includes URLs, search form URLs, description of domains, and vendor descriptions, wherein the vendor descriptions include generalized rules about how product information is organized on each of the vendor sites;
- b. processing parameters for a price comparison request for a desired product using the information maintained in the offline database, while price comparison request is received from an online user;
- c. extracting real-time price and product information from identified ones of the plurality of vendor sites, wherein the extracted price and product information are in a native language of the site; and
- d. displaying the extracted price and product information to the user.

30. The method of claim 29, wherein the step of receiving a price comparison request further comprises the step of receiving from the user at least one search parameter and an identification of at least one online vendor on said computer network; and further wherein the extracting step includes the steps of

- i. posting a request using the processed parameters to at least one of the plurality of vendors online, in real-time; and
- ii. retrieving data related to the price and product information from search results obtained in response to the posting step;
- iii. sorting the retrieved data by price; and
- iv. displaying processed data for the desired product from at least one of the plurality of vendors.

31. The method of claim 30, wherein the step of receiving a price comparison request is initiated by the online user.

32. The method of claim 30, wherein in the step of posting a request, the processed parameters are the combination of the search parameters and vendor identification received from the user, vendor description for the identified vendor, and the URL of the identified vendor.

33. The method of claim 30, wherein the vendor descriptions maintained in the offline database includes patterns which identify information in vendor sites, and further wherein the step of retrieving data employs the patterns.

34. The method of claim 30, wherein the step of extracting real-time price and product information is domain-independent and language-independent.

35. The method of claim 30, wherein the step of displaying the processed data is based on composing information to be displayed in HTML.